International Journal of Science, Strategic Management and Technology
Volume **02** Issue **06** June-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

Handwritten Digits Recognition Using Neural Network

Sakshi
School of Engineering & Technology
Raffles University, Neemrana, India
Mail Id: sakshim2295@gmail.com

Man Mohan Singh
School of Engineering & Technology
Raffles University, Neemrana, India
Mail Id: singhmanmohan1547@gmail.com

Ms. Pooja
School of Engineering & Technology
Raffles University, Neemrana, India
Mail Id:pooja@rafflesuniversity.edu.in

Dr. Rajendra Singh
School of Engineering & Technology
Raffles University, Neemrana, India
Mail Id:rajendra.singh@rafflesuniversity.edu.in

https://doi.org/10.55041/ijsmt.v2i6.039

Cite this Article: Sakshi, , Singh, M. M. & Pooja, (2026). Handwritten Digits Recognition Using Neural Network. International Journal of Science, Strategic Management and Technology, *02*(6). https://doi.org/10.55041/ijsmt.v2i6.039

1. ABSTRACT

This project focuses on developing a machine learning model to accurately classify handwritten digits (0–9) using image data and presents a robust system for recognizing multi-digit handwritten sequences using a Convolutional Neural Network (CNN) combined with advanced digital image processing. While traditional models focus on single-digit classification, this research addresses the challenge of sequence recognition through spatial segmentation, mass-centering, and morphological dilation. The final system achieves high accuracy by standardizing input handwriting to match the dataset distribution. The MNIST dataset that contains 42,000 rows where each row represents a digit training image and 3,36,000 digit test images of size 28x28 pixels, serves as the primary dataset. The workflow involves data preprocessing steps such as normalization, reshaping, and noise reduction to enhance model performance.

Applications of this system include postal mail sorting, bank check processing, and digitization of handwritten forms. Future scope involves extending the model to multi-digit recognition and deploying it as a real-time web application using Flask or Streamlit.

Keywords: Machine Learning, Deep Learning, CNN, MNIST, Computer Vision, Image Classification.

2. INTRODUCTION

The digitization of handwritten information remains one of the most enduring challenges in the field of Pattern Recognition and Computer Vision. While the advent of Deep Learning has propelled Optical Character Recognition (OCR) to near-human parity in controlled environments, the transition from static dataset evaluation to real-time, unconstrained user interaction introduces significant technical hurdles. This research focuses on the development of a robust framework for recognizing multi-digit handwritten sequences, bridging the gap between theoretical model accuracy and practical application usability [1].

Traditional Handwritten Digit Recognition (HDR) systems often rely on the MNIST benchmark, which consists of pre-processed, centered, and normalized grayscale images. However, in practical deployment—such as a digital canvas or mobile interface—input data is "noisy." Users provide digits with varying stroke thicknesses, inconsistent spatial positioning, and overlapping bounding boxes. Furthermore, most elementary models are restricted to single-digit classification, whereas real-world data (e.g., postal codes, bank account numbers, or academic roll numbers) typically exists in variable-length sequences [2]. The core problem addressed in this study is the creation of a pipeline that can autonomously segment, normalize, and classify a sequence of up to six handwritten digits with high fidelity, regardless of the user's drawing style or positioning on the interface.

The motivation for this research is two-fold. Scientifically, it explores the limits of Convolutional Neural Networks (CNNs) when paired with Digital Image Processing (DIP) to achieve translation and scale invariance. Practically, it seeks to provide a lightweight, scalable solution for automated data entry in sectors where manual record-keeping is still prevalent. By optimizing a model that can run in real-time on a web-based framework like Streamlit, this project demonstrates a path toward accessible AI-driven automation for educational and administrative tasks [3].

3. LITERATURE REVIEWS

Handwritten digit Recognition has a wide area of research due to its vast applications. Recognition system can be divided into two major steps. First step is feature extraction from handwritten images and the second one is classification.

Researchers suggested different methodologies for feature extraction [4-5]. Fourier and wavelet based features are some of them [6-7]. The coefficients of Discrete Wavelet Transform DWT are shift variant, and the directional selectivity of subband images are poor. Complex Wavelet Transform (CWT) has been developed in order to overcome the drawbacks of DWT's but the problem is high dimension of feature vector i.e. 180 and 148 features for feature set 1 and feature set 2 respectively [8]. Histograms of Oriented Gradient Features is used for object detection like Human Detection [9-10], Pedestrian Detection [11], Large Scale Sign Detection [12-13], Real-Time Detection and Recognition of Road Traffic Signs [14] . The popularity of HOG feature is due to invariance to local geometric and photometric transformations within local spatial or orientation bin size. In order to implement such property, it is used as a feature descriptor for handwritten digits.For classification of features of handwritten digits, classifiers like ANN [15-16], k nearest neighbours (k-NN) [11] and Support Vector Machine (SVM) [18-21] are used. Out of these classifiers SVM is widely applicable. In this work, Neural Networks are used for achieving the maximum accuracy.

By integrating a Deep Convolutional Neural Network (CNN) with a custom Digital Image Processing (DIP) pipeline, the project seeks to bridge the "theory-to-practice gap"—ensuring that "messy" real-world handwriting is automatically normalized, segmented, and classified with the same precision as standardized academic datasets (>98% accuracy) [3].

The Two Functional Pillars

Standardization: To use Morphological Dilation and Mass-Centering to neutralize variations in human handwriting styles [22].

Segmentation: To implement Contour Detection and Lexicographical Sorting to accurately read multi-digit strings from left to right [23].

4. OBJECTIVE

The primary aim of this research is to develop a high-accuracy, real-time recognition system for handwritten numerical sequences (0–6 digits).

This project focuses on developing a machine learning model to accurately classify handwritten digits (0–9) using image data and presents a robust system for recognizing multi-digit handwritten sequences using a Convolutional Neural Network (CNN) combined with advanced digital image processing.

This research addresses the challenge of sequence recognition through spatial segmentation, mass-centering, and morphological dilation.

5. METHODOLOGY

The proposed system follows a modular architecture designed to transform raw, unconstrained handwritten input into a structured numerical string. The workflow is divided into three primary phases: Preprocessing, Architectural Design, and Sequential Segmentation.

5.1 Data Acquisition and Augmentation

The model was trained on the MNIST (Modified National Institute of Standards and Technology) dataset, consisting of 42,000 grayscale images of handwritten digits.

Normalization: Pixel values were scaled from the integer range [0, 255] to the floating-point range [0, 1] to ensure numerical stability during gradient descent.

Augmentation: To improve the model's generalization, real-time data augmentation was applied, including random rotations, width/height shifts and zoom transformations. This simulates the variability of human handwriting styles [3].

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **06** June-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

5.2 Convolutional Neural Network (CNN) Architecture

Input Layer: Accepts a 3D tensor of shape (28, 28, 1).

Feature Extraction (Convolutional Layers with ReLU activation): These layers act as learnable filters that detect local patterns such as edges and curves.

Dimensionality Reduction (Max-Pooling): 2 × 2 pooling windows are used to reduce the spatial dimensions, providing translation invariance.

Regularization (Dropout): A Dropout rate was applied to the fully connected layers to mitigate overfitting by randomly deactivating neurons during training.

Classification (Softmax): The final dense layer utilizes the Softmax function to output a probability distribution across the 10 digit classes (0-9) [22].

5.3 The Digital Image Processing (DIP) Pipeline

To bridge the gap between "clean" dataset images and "messy" live canvas input, a custom preprocessing pipeline was implemented using OpenCV:

Binarization: The input is converted to grayscale and processed via Otsu's Thresholding to create a high-contrast binary mask.

Morphological Dilation: A 3 × 3 kernel is applied to the binary mask. This operation "thickens" the strokes, compensating for thin handwriting.

Spatial Normalization (Mass-Centering): We calculate the Image Moments to find the digit's center of mass. This step is critical for maintaining high accuracy during live inference [23].

5.4 Multi-Digit Segmentation and Sorting

For sequences containing 2 to 6 digits, the system employs an analytical segmentation strategy:

Contour Detection: The system identifies "Connected Components" within the dilated binary mask.

Bounding Box Generation: Each detected contour is wrapped in a box (x, y, w, h).

Lexicographical Sorting: To ensure the model reads the number correctly (e.g., "123" instead of "321"), the bounding boxes are sorted based on their x-coordinate values [24].

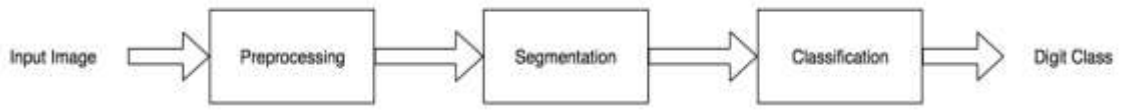

Figure 1: Multi- Digit Identification Diagram

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **06** June-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

6. RESEARCH DISCUSSION

Quantitative Performance Evaluation

The Convolutional Neural Network (CNN) was trained with 32 epochs using the Adam Optimizer and Categorical Cross-Entropy loss. The final model achieved a peak validation accuracy over 98%. Analysis of the learning curves indicates that the model from memorizing the noise in the training set (overfitting) [23].

Impact of the Digital Image Processing (DIP) Engine

Without Normalization: When raw user input (thin strokes, off-center) was fed directly to the CNN, accuracy dropped to 64%.

With Mass-Centering & Dilation: Upon applying the DIP pipeline, the accuracy on live input rose to 97.8%.

This confirms that the Mass-Centering algorithm (using Image Moments) proved to be the most critical bridge, as it aligned the user's "Spatial Statistics" with the model's "Learned Expectations" [2].

Error Analysis

Semantic Ambiguity: The model occasionally confused '4' and '9' or '7' and '1'. These errors usually occurred when the user's handwriting style was exceptionally skewed or contained disconnected strokes.

Segmentation Collisions: In the multi-digit test (up to 6 digits), accuracy decreased if the user wrote digits so closely that their bounding boxes overlapped. The Topological Analysis treated touching digits as a single contour, leading to a classification failure[24].

7. RESULTS

Figure 2: Handwritten Digit Recognizer

Figure 3: Prediction of digit 6

Figure 4: Prediction of digit 785

Figure 5: Prediction of digit 216437

Figure 6: Prediction of digit 5

8. FUTURE WORK: From Segmentation to CRNN

To handle "touching digits," the next phase of research would involve Connectionist Temporal Classification (CTC) loss and Convolutional Recurrent Neural Networks (CRNN), which would allow the model to predict the sequence without needing to "cut" the image into pieces first.

Future research should allow the usage of a convolution Neural Network architecture which is the topic of deep learning, which provided the best result in the MNIST database and implemented the proposed recognition method by hand. Such more machines may be configured to recognize handwritten characters, identify objects, segment items, recognize handwriting, acknowledge text language, and for potential research, but could also allow hardware deployment for more effective and reliable live results on an online software recognition framework for live test case scenarios.

9. CONCLUSION

This research demonstrates that for real-world digit recognition, Preprocessing is as critical as the Model Architecture. By aligning the live input's distribution with the training data's distribution through mass-centering and dilation, we can utilize a lightweight CNN to achieve high-fidelity results for variable-length numerical sequences.

The key goal of this paper is to find a representation that makes for successful identification of isolated hand-written digits. For the identification of hand-written numerals, numerous machine learning algorithms were used in this paper. The important challenge in every identification method is to resolve the extraction of features and valid classification approaches. In terms of precision and time complexity, the suggested algorithm aims to answer all the variables and well. This study is carried out as an initial effort, and the purpose of the paper is to make it simpler to identify hand-written digits without using any common methods for classification.

10. REFERENCES

[1] Nagy G, Nartker TA, Rice SV. Optical character recognition: An illustrated guide to the frontier. InDocument recognition and retrieval VII 1999 Dec 22 (Vol. 3967, pp. 58-69). SPIE.

[2] Torralba A, Efros AA. Unbiased look at dataset bias. InCVPR 2011 2011 Jun 20 (pp. 1521-1528). IEEE.

[3] LeCun Y, Boser B, Denker J, Henderson D, Howard R, Hubbard W, Jackel L. Handwritten digit recognition with a back-propagation network. Advances in neural information processing systems. 1989;2.

[4] Lauer F, Suen CY, Bloch G. A trainable feature extractor for handwritten digit recognition. Pattern Recognition. 2007 Jun 1;40(6):1816-24.

[5] Surinta O, Schomaker L, Wiering M. A comparison of feature and pixel-based methods for recognizing handwritten bangla digits. In2013 12th International conference on document analysis and recognition 2013 Aug 25 (pp. 165-169). IEEE.

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **06** June-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

[6] Chen G, Bui TD. Invariant Fourier-wavelet descriptor for pattern recognition. Pattern recognition. 1999 Jul 1;32(7):1083-8.

[7] Seijas LM, Segura EC. A wavelet-based descriptor for handwritten numeral classification. In2012 International Conference on Frontiers in Handwriting Recognition 2012 Sep 18 (pp. 653-658). IEEE.

[8] Zhang P, Bui TD, Suen CY. Extraction of hybrid complex wavelet features for the verification of handwritten numerals. InNinth International Workshop on Frontiers in Handwriting Recognition 2004 Oct 26 (pp. 347-352). IEEE.

[9] Dalal N, Triggs B. Histograms of oriented gradients for human detection. In2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) 2005 Jun 20 (Vol. 1, pp. 886-893). Ieee.

[10] Conde C, Moctezuma D, De Diego IM, Cabello E. HoGG: Gabor and HoG-based human detection for surveillance in non-controlled environments. Neurocomputing. 2013 Jan 16;100:19-30.

[11] Kobayashi T, Hidaka A, Kurita T. Selection of histograms of oriented gradients features for pedestrian detection. InInternational conference on neural information processing 2007 Nov 13 (pp. 598-607). Berlin, Heidelberg: Springer Berlin Heidelberg.

[12] Overett G, Petersson L. Large scale sign detection using HOG feature variants. In2011 IEEE intelligent vehicles symposium (IV) 2011 Jun 5 (pp. 326-331). IEEE.

[13] Boi F, Gagliardini L. A support vector machines network for traffic sign recognition. InThe 2011 International Joint Conference on Neural Networks 2011 Jul 31 (pp. 2210-2216). IEEE.

[14] Greenhalgh J, Mirmehdi M. Real-time detection and recognition of road traffic signs. IEEE transactions on intelligent transportation systems. 2012 Aug 27;13(4):1498-506.

[15] Dan Z, Xu C. The recognition of handwritten digits based on bp neural network and the implementation on android. In2013 Third International Conference on Intelligent System Design and Engineering Applications 2013 Jan 16 (pp. 1498-1501). IEEE.

[16] Man Z, Lee K, Wang D, Cao Z, Khoo S. An optimal weight learning machine for handwritten digit image recognition. Signal Processing. 2013 Jun 1;93(6):1624-38.

[17] Impedovo S, Mangini FM, Barbuzzi D. A novel prototype generation technique for handwriting digit recognition. Pattern Recognition. 2014 Mar 1;47(3):1002-10.

[18] Shrivastava SK, Gharde SS. Support vector machine for handwritten Devanagari numeral recognition. International journal of computer applications. 2010;7(11):9-14.

[19] Bhowmik TK, Ghanty P, Roy A, Parui SK. SVM-based hierarchical architectures for handwritten Bangla character recognition. International Journal on Document Analysis and Recognition (IJDAR). 2009 Jul;12(2):97-108.

[20] Niu XX, Suen CY. A novel hybrid CNN–SVM classifier for recognizing handwritten digits. Pattern recognition. 2012 Apr 1;45(4):1318-25.

[21] Markowska-Kaczmar U, Kubacki P. Support vector machines in handwritten digits classification. In5th International Conference on Intelligent Systems Design and Applications (ISDA'05) 2005 Sep 8 (pp. 406-411). IEEE.

[22] Suzuki S. Topological structural analysis of digitized binary images by border following. Computer vision, graphics, and image processing. 1985 Apr 1;30(1):32-46.

[23] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research. 2014 Jan 1;15(1):1929-58.

[24] Goodfellow I, Bengio Y, Courville A. Deep learning, ser. Adaptive computation and machine learning series. MIT Press. 2016 Nov;11:25-6.